



Europäisches Patentamt  
European Patent Office  
Office européen des brevets



(11) **EP 0 990 996 A2**

(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:  
05.04.2000 Bulletin 2000/14

(51) Int. Cl.<sup>7</sup>: **G06F 17/30**

(21) Application number: **99203040.3**

(22) Date of filing: **17.09.1999**

(84) Designated Contracting States:  
**AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU  
MC NL PT SE**  
Designated Extension States:  
**AL LT LV MK RO SI**

(30) Priority: **30.09.1998 US 163618**

(71) Applicant: **EASTMAN KODAK COMPANY**  
Rochester, New York 14650 (US)

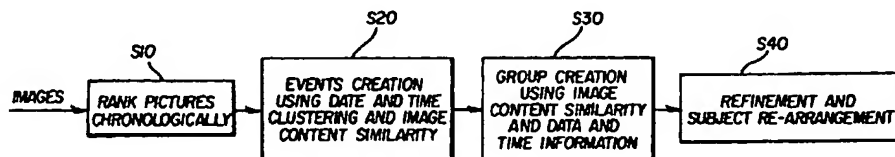
(72) Inventors:  
• **Loul, Alexander Cheng Pong,**  
Eastman Kodak Company  
Rochester, New York 14650-2201 (US)  
• **Pavie, Eric Serge,**  
Eastman Kodak Company  
Rochester, New York 14650-2201 (US)

(74) Representative:  
**Lewandowsky, Klaus, Dipl.-Ing. et al**  
Kodak Aktiengesellschaft,  
Patentabteilung  
70323 Stuttgart (DE)

(54) **A method for automatically classifying images into events**

(57) A method for automatically classifying images into events, the method includes the steps of: receiving a plurality of images having either or both date and/or time of image capture; determining one or more largest time differences of the plurality of images based on clustering of the images; and separating the plurality of

images into the events based on having one or more boundaries between events which one or more boundaries correspond to the one or more largest time differences.



**FIG. 1**

**EP 0 990 996 A2**

**Description****FIELD OF THE INVENTION**

[0001] The invention relates generally to the field of 5  
image processing having image understanding that  
automatically classifies pictures by events and the like  
and, more particularly, to such automatic classification  
of pictures by time and date analysis and by block-  
based analysis which selectively compares blocks of 10  
the images with each other.

**BACKGROUND OF THE INVENTION**

[0002] Pictorial images are often classified by the par- 15  
ticular event, subject or the like for convenience of  
retrieving, reviewing, and albuming of the images. Typi-  
cally, this has been achieved by manually segmenting  
the images, or by the below-described automated  
method. The automated method includes grouping by 20  
color, shape or texture of the images for partitioning the  
images into groups of similar image characteristics.

[0003] Although the presently known and utilized 25  
methods for partitioning images are satisfactory, there  
are drawbacks. The manual classification is obviously  
time consuming, and the automated process, although  
theoretically classifying the images into events, is sus-  
ceptible to miss-classification due to the inherent inac-  
curacies involved with classification by color, shape or  
texture. 30

[0004] Consequently, a need exists for overcoming the  
above-described drawbacks.

**SUMMARY OF THE INVENTION**

[0005] The present invention is directed to overcoming 35  
one or more of the problems set forth above. Briefly  
summarized, according to one aspect of the present  
invention, the invention resides in a method for automati-  
cally classifying images into events, the method com-  
prising the steps of: receiving a plurality of images  
having either or both date and/or time of image capture;  
determining one or more largest time differences of the  
plurality of images based on clustering of or more larg- 40  
est time differences of the plurality of images based on  
clustering of the images; and separating the plurality of  
images into the events based on having one or more  
boundaries between events which one or more bounda-  
ries correspond to the one or more largest time differ-  
ences.

[0006] These and other aspects, objects, features and 45  
advantages of the present invention will be more clearly  
understood and appreciated from a review of the follow-  
ing detailed description of the preferred embodiments  
and appended claims, and by reference to the accom-  
panying drawings. 50

**ADVANTAGEOUS EFFECT OF THE INVENTION**

[0007] The present invention has the advantage of  
improved classification of images by utilizing both date  
and time information and block-based comparison that  
checks for similarity of subject and background in the  
images. If date and time information is not available,  
then the block-based analysis may be used as the sole  
basis for classification.

**BRIEF DESCRIPTION OF THE DRAWINGS**

[0008]

Fig. 1 is a block diagram illustrating an overview of  
the present invention;

Fig. 2 is a block diagram illustrating a date and time  
clustering technique of the present invention;

Fig. 3 is a graph illustrating a scaling function used  
to map the result of the 2-means clustering;

Fig. 4 is a graph illustrating a typical result of the  
scaling function of Fig. 3;

Fig. 5 is a diagram illustrating a block diagram of an  
event boundary checking after the date and time  
clustering;

Fig. 6 is a diagram illustrating grouping of images  
within each event based on content;

Fig. 7 is a block diagram of a group-merging step of  
the present invention;

Fig. 8 is a block diagram of image re-arrangement  
within each group;

Fig. 9 is a flowchart of block-based histogram corre-  
lation technique;

Fig. 10 is diagram illustrating the comparison  
between block histogram;

Fig. 11 is diagram of an example of best intersec-  
tion mapping for three segment analysis; and,

Fig. 12 is an illustration of shift detection within the  
block based histogram correlation.

**DETAILED DESCRIPTION OF THE INVENTION**

[0009] In the following description, the present inven-  
tion will be described in the preferred embodiment as a  
software program. Those skilled in the art will readily  
recognize that the equivalent of such software may also  
be constructed in hardware.

[0010] Still further, as used herein, computer readable  
storage medium may comprise, for example; magnetic  
storage media such as a magnetic disk (such as a  
floppy disk) or magnetic tape; optical storage media  
such as an optical disc, optical tape, or machine reada-  
ble bar code; solid state electronic storage devices such  
as random access memory (RAM), or read only mem-  
ory (ROM); or any other physical device or medium  
employed to store a computer program.

[0011] In addition, the term "event" is defined herein  
as a significant occurrence or happening as perceived

by the subjective intent of the user of the image capture device.

[0012] Before describing the present invention, it facilitates understanding to note that the present invention is preferably utilized on any well-known computer system, such as a personal computer. Consequently, the computer system will not be discussed in detail herein. It is also instructive to note that the images are either directly input into the computer system (for example by a digital camera) or digitized before input into the computer system (for example by scanning).

[0013] Referring to now Fig. 1, there is illustrated a flow diagram illustrating an overview of the present invention. Digitized images are input into the computer system where a software program of the present invention will classify them into distinct categories. The images will first be ranked S10 in chronological order by analyzing the time and date of capture of each image. The date and/or time of capture of each picture may be extracted, for example, from the encoded information on the film strip of the Advanced Photo System (APS) images, or from information available from some digital cameras. The representations of the images will then be placed S20 into one of a plurality of distinct events by a date and time clustering analysis that is described below. Within each event, the contents of the images are analyzed S20 for determining whether images closest in time to an adjacent event should be maintained in the event as defined by the clustering analysis, or the adjacent events merged together. After the images are defined into events, a further sub-classification (grouping) within each event is performed. In this regard, the images within each event will then be analyzed by content S30 for grouping images of similar content together, and then the date and time S30 for further refining the grouping.

[0014] The event segmentation S20 using the date and time information is by a k-means clustering technique, as will be described in detail below, which groups the images into events or segments. A boundary check is then performed on the segments S20 for verifying that the boundary images should actually be grouped into the segment identified by the clustering, as will also be described below.

[0015] These groups of images are then sent to a block-based histogram correlator S30 for analyzing the content. For each event or segment sent to the correlator, a content or subject grouping S30 is performed thereon for further sub-classifying the images by subject within the particular event segment. For example, within one event, several different subjects may appear, and these subject groupings define these particular subjects. The subject grouping is based primarily on image content, which is performed by a block-based histogram correlation technique. This correlation compares portions of two images with each other, as will also be described in detail below. The result of the ranking is the classification of images of each segment into distinct

subject groupings. The date and time of all the images within each subject grouping are then compared to check whether any two or more subject grouping can be merged into a single subject grouping S30.

[0016] A refinement and subject re-arrangement analysis S40 will further improve the overall classification and the subject grouping by rearranging certain images within a subject group.

[0017] Referring to Fig. 2, there is shown an exploded block diagram illustrating the data and time clustering technique S20. First, the time interval between adjacent pictures (time difference) is computed S20a. A histogram of the time differences is then computed S20b, an example of which is shown in block 10. The abscissa of the histogram is preferably the time in minutes, and the ordinate of the histogram is the number of pictures having the specified time difference. The histogram is then mapped S20c to a scaled histogram using a time difference scaling function, which is shown in Fig. 3. This mapping is to take the pictures with small time difference and substantially maintain its time difference, and to take pictures with a large time difference and compress their time difference.

[0018] A 2-means clustering is then performed S20d on the mapped time-difference histogram for separating the mapped histogram 10 into two clusters based on the time difference; the dashed line represents the separation point for the two clusters. For further details of 2-means clustering, Introduction to Statistical Pattern Recognition, 2<sup>nd</sup> edition by Keinosuke Fukunaga 1990 may be consulted, and therefore, the process of 2-means clustering will not be discussed in detail herein. Referring briefly to Fig. 4, the result of 2-means clustering is the segmentation of the histogram into two portions 10a and 10b. Normally, events are separated by large time differences. The 2-means clustering, therefore, is to define where these large time differences actually exist. In this regard, the right portion 10b of the 2-means clustering output defines the large time differences that correspond to the event boundaries.

[0019] Referring to Fig. 5, there is illustrated an example of boundary checking between events. For two consecutive events i and j, a plurality of block-based, histogram comparisons are made to check if the pictures at the border of one event are different from the pictures at the border of the other event. If the comparison of content is similar, the two segments are merged into one segment. Otherwise, the segments are not merged. Preferably, the comparisons are performed on the three border images of each event (i3, i4, i5 with j1, j2, j3), as illustrated in the drawing. For example, image i5 is compared with image j1 and etc. This block-based histogram comparison technique will be described in detail hereinbelow.

[0020] Referring to Fig. 6, there is illustrated an overview of subject (content) grouping for each segmented event. Within each segmented event i, adjacent pictures are compared (as illustrated by the arrows) with each

other using the below-described, block-based histogram technique. For example, the block-based histogram technique may produce five subject groupings (for example groups 1-5) from the one event  $i$ , as illustrated in the drawing. The arrangement of the subject grouping is stored for future retrieval during the subject arrangement step S40. In particular, the subject grouping having a single image is stored (for example groups 2, 3, and 5).

[0021] Referring to Fig. 7, after the grouping by content, a time and date ordering is performed on the groupings for merging groups together based on a time and date analysis. A histogram of the time difference between adjacent images in the event is computed, similar to Fig. 4. A predetermined number of the largest time differences (for example boundary  $a_{12}$ ) are compared with the boundaries (for example boundaries  $b_{12}$ ,  $b_{23}$ ,  $b_{34}$ ,  $b_{45}$ ) of the subject grouping determined by the block-based analysis. The predetermined number of largest time differences are determined by dividing the total number of images within an event by the average number of picture per group (four is used in the present invention). If the boundary of the subject grouping matches the boundary based on the chosen time differences, the subject groupings will not be merged. If there is not a match between the two boundaries, the subject groupings having a boundary not having a matched time difference in the histogram will be merged into a single subject grouping (for example groups  $b_1$ ,  $b_6$ ,  $b_3$  merged into resulting group  $c_1$ ).

[0022] Referring to Fig. 8, there is illustrated a diagram of image re-arrangement within each group. The arrangement of the initial subject groupings is retrieved for identifying subject groupings that contain single images (for example the groups with a single image of Fig. 6 -groups 2, 3, and 5 that are re-illustrated as groups 2, 3, and 5 in Fig. 8). Any single images from the same subject grouping that are merged as identified by the merged subject grouping (for example, groups  $c_1$  and  $c_2$  of Fig. 7) are compared with all other images in the merged subject grouping, as illustrated by the arrows. This comparison is based on block-based histogram analysis. If the comparisons are similar, these images will be re-arranged so that the similar images are located adjacent each other, for example groups  $d_1$  and  $d_2$ .

[0023] Further refinement is done by comparing any group that still contains a single image after the above procedure, with all the images in the event. This is to check if these single image groups can be better arranged within the event grouping. This comparison is similar to the subject re-arrangement step of Fig. 8.

[0024] Referring to Fig. 9, there is illustrated a flow-chart of the block-based histogram correlation used in the above analyses. First, a histogram of the entire image of both images is computed S50, a global histogram. A comparison of the two histograms is performed by histogram intersection value S60 illustrated the fol-

lowing equation:

$$Inter(R, C) = \frac{\sum_{i=1}^n \min(R_i, C_i)}{\sum_{i=1}^n R_i}$$

where  $R$  is the histogram of the reference image,  $C$  is the histogram of the candidate image, and  $n$  is the number of bins in the histogram. If the intersection is under a threshold S65, preferably 0.34, although other thresholds may be used, the images are different. If the threshold is met or exceeded S65, then a block-based histogram correlation will be performed S70. In this regard, each image will be divided into blocks of a given size, preferably 32 x 32 pixels in the present invention. It is instructive to note that those skilled in the art may vary the block size depending on the resolution of the image without departing from the scope of the invention. For each block, a color histogram is computed. Referring to Fig. 10, if one image is considered a reference image and one image a candidate image, the images are compared in the following way. Each block 20 of the reference image is compared to the corresponding block 30 of the candidate image and to the adjacent blocks 40, 8 blocks in the present invention.

[0025] Referring to Fig. 9, the block histograms between the reference image and the candidate image are compared using the histogram intersection equation defined above S80. The average intersection value is derived by computing the average of the best intersection values from each of the block comparisons S90. This average intersection value will be compared to a low threshold (preferably 0.355), and a high threshold (preferably 0.557). If the average intersection value is below the low threshold S95, the two images are considered different. If the average intersection value is above the high threshold S96, then the two images are considered similar. If the average intersection value is between these two thresholds, further analysis will be performed as described below (3-segment average intersection map S100).

[0026] Referring to both Figs. 9 and 11, a 3-segment analysis will be performed to determine if the two images may contain a similar subject. This is performed by first forming a map 60 which contains the average of the two highest intersection values of each of the block comparisons; for example, 9 comparisons were performed in the illustration of Fig. 10, the average of the highest two will be used for map 60. Fig. 11 illustrates, for example, a 9 x 6 block although it should be understood that the map size depends on the size of the image. This map is divided into three parts: the left portion 70a, the center portion 70b, and the right portion 70c. If the average intersection value of the center portion 70b is higher than a threshold (preferably 0.38)

S105, the two images may contain a very similar subject in the center portion 70b of the image, and the two images may be considered to be similar by subject. In addition, the comparisons of the histogram will be performed with the reference and candidate images reversed. If the two images are similar both methods should give substantially similar correlation; obviously if they are different, the results will not be similar. The images are then checked S110 to determine if there is a high intersection value in one of the directions, right, left, up, and down.

[0027] Referring to Figs. 9 and 12, shift detection is used to determine the case when the two images 90 and 100 (of two different sizes in the drawing) have very similar subject that appears in different locations of the image. For example, the main subject may be situated in the center of one image and to the left-hand side of the other image. Such a shift can be determined by recording both the best intersection values of the reference blocks, as well as the coordinates of the corresponding candidate blocks. This is achieved by comparing the intersection values of the blocks in four directions (right, left, up, and down). The entire image will be shifted by one block (as illustrated by the arrows) in one of the directions (right in the drawing) where the best intersection value is the highest. The above analysis and the shift can be repeated S120 to check for similarity.

[0028] The invention has been described with reference to a preferred embodiment. However, it will be appreciated that variations and modifications can be effected by a person of ordinary skill in the art without departing from the scope of the invention.

#### Claims

1. A method for automatically classifying images into events, the method comprising the steps of:

- (a) receiving a plurality of images having either or both date and/or time of image capture;
- (b) determining one or more largest time differences of the plurality of images based on clustering of the images; and,
- (c) separating the plurality of images into the events based on having one or more boundaries between events which one or more boundaries correspond to the one or more largest time differences.

2. The method as in claim 1, wherein step (b) includes computing a time difference histogram and performing a 2-means clustering on the time difference histogram for defining the one or more boundaries.

3. The method as in claim 2, wherein step (b) further includes mapping the time difference histogram through a time difference scaling function before

performing the 2-means clustering.

4. The method as in claim 2, wherein step (c) includes checking the images adjacent the one or more boundaries for similarity by comparing content of the images.
5. The method as in claim 4, wherein step (c) includes checking the images adjacent the one or more boundaries for similarity by using a block-based histogram correlation technique.
6. The method as in claim 5 further comprising step (d) dividing the events into subject grouping by using an image content analysis.
7. The method as in claim 6, wherein step (d) includes dividing the events into subject grouping by using a block-based histogram technique.
8. A method for automatically classifying images into events, the method comprising the steps of:

- (a) receiving a plurality of images arranged in chronological order;
- (b) dividing the images into a plurality of blocks; and,
- (c) grouping the images into subject grouping based on block-based histogram correlation which includes computing a color histogram of each block and computing a histogram intersection value which determines the similarity between blocks.

9. The method as in claim 8, wherein step (c) includes comparisons of two of the images by shifting one of the images in a desired direction based on the intersection value and then computing the block based correlation.
10. The method as in claim 9, wherein step (c) includes forming a map that contains two best intersection values of each of the block comparisons; dividing the map into three portions; and then comparing center portions for similarity.

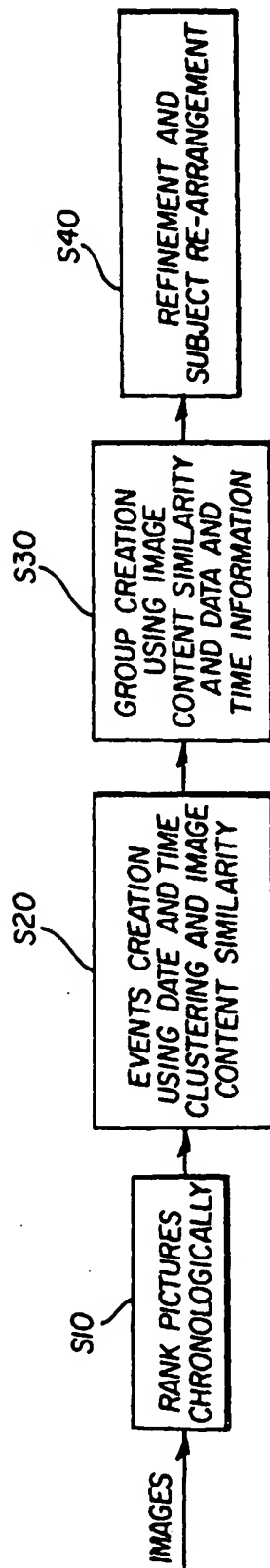


FIG. 1

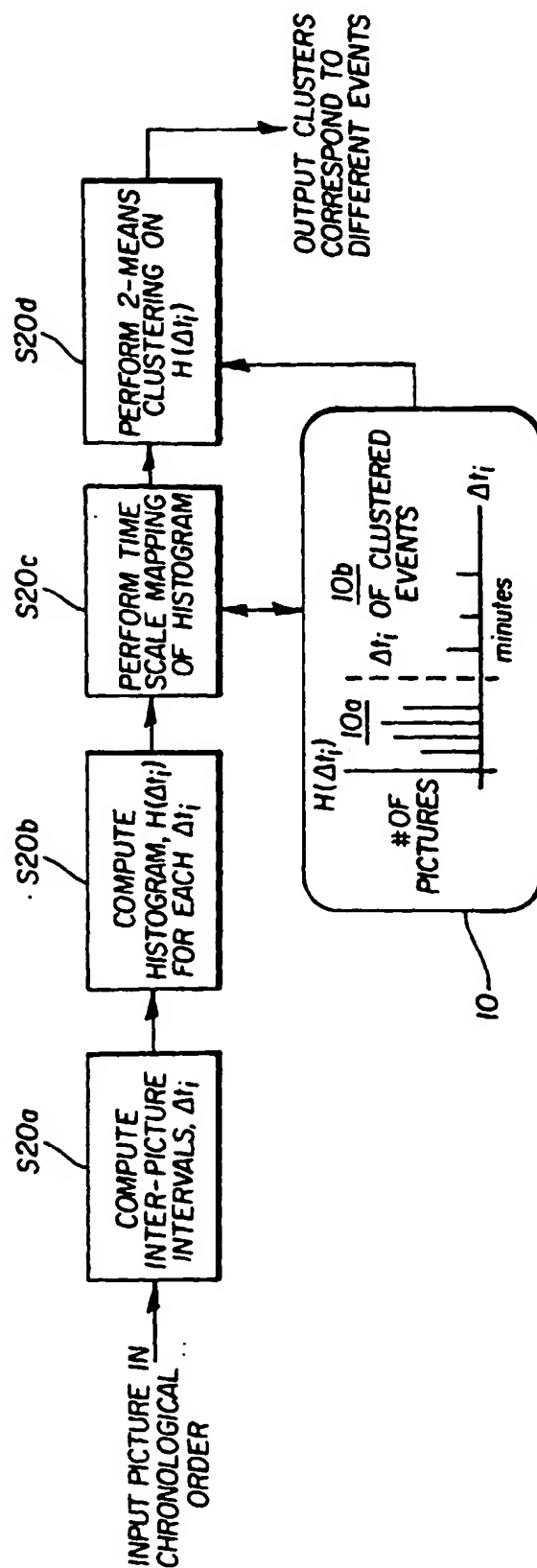


FIG. 2

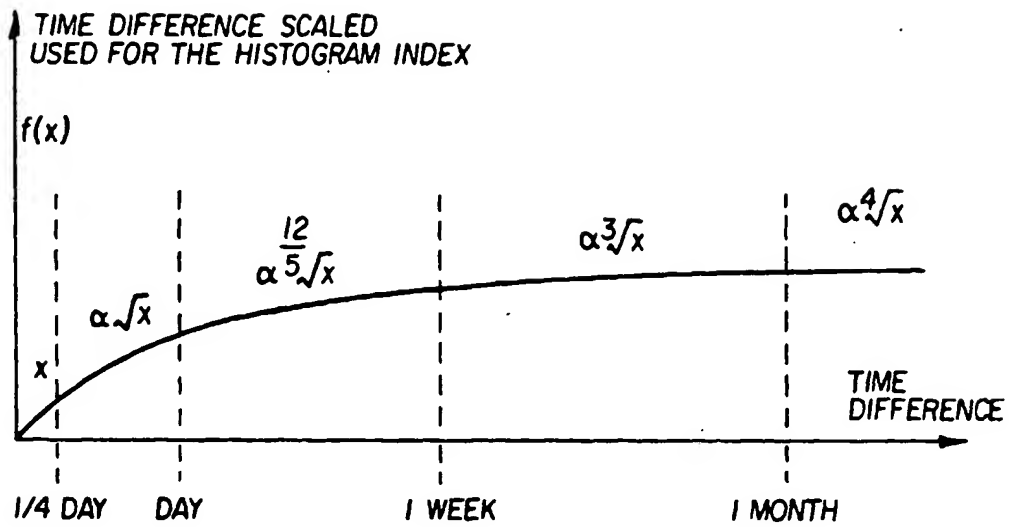


FIG. 3

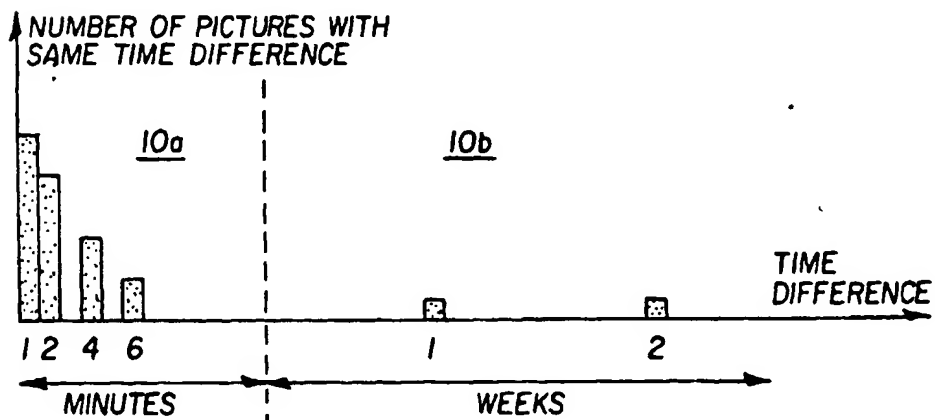


FIG. 4



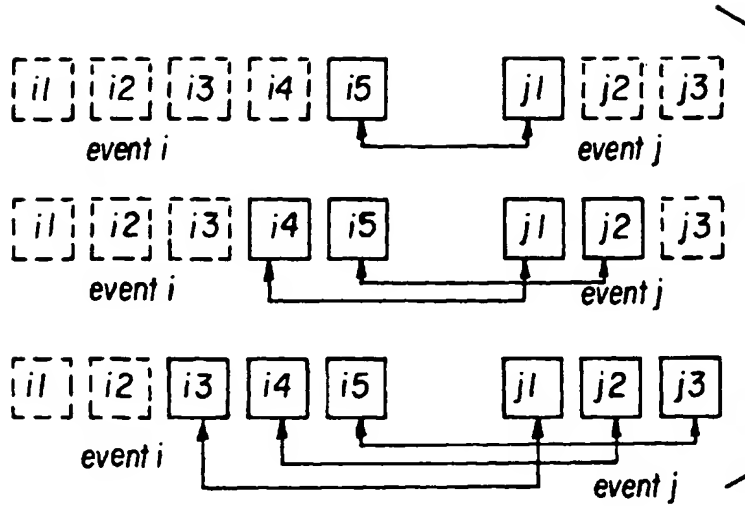


FIG. 5

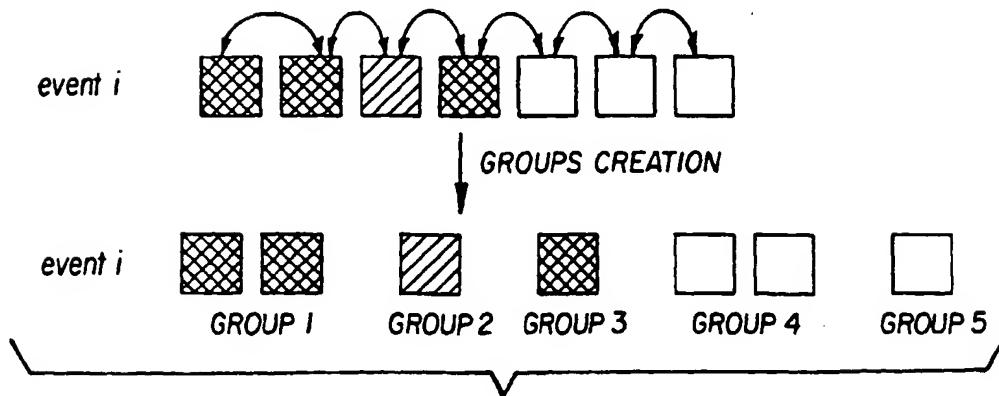


FIG. 6

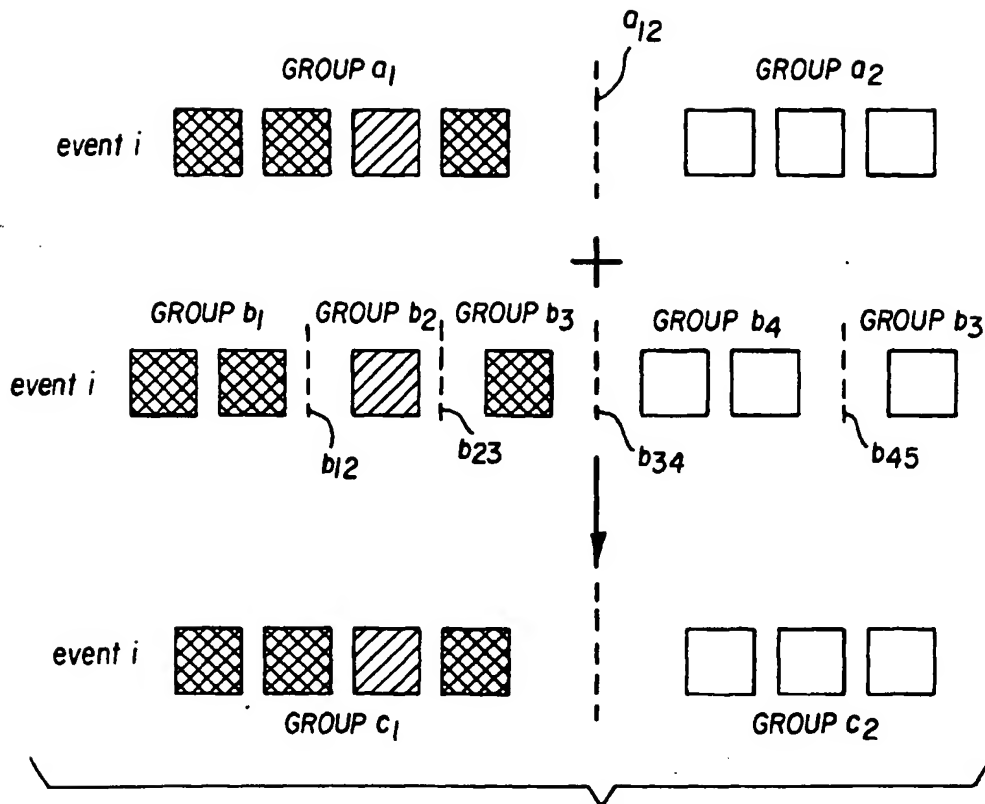


FIG. 7

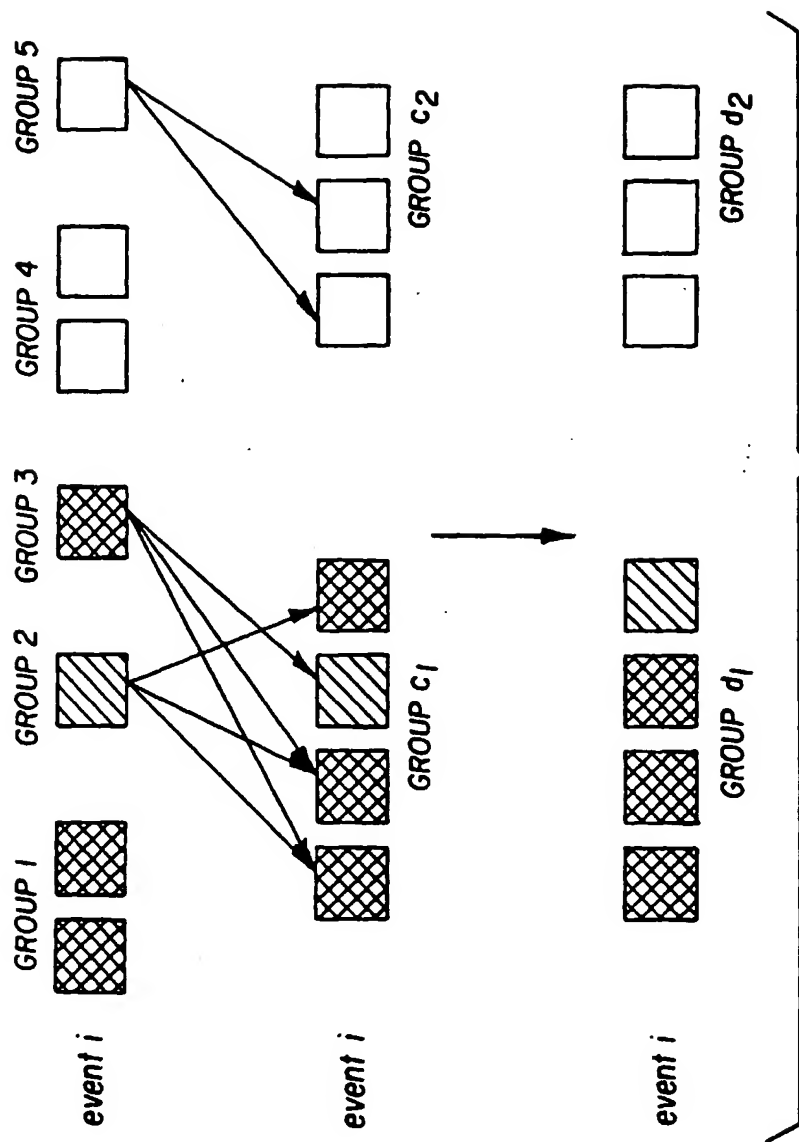


FIG. 8

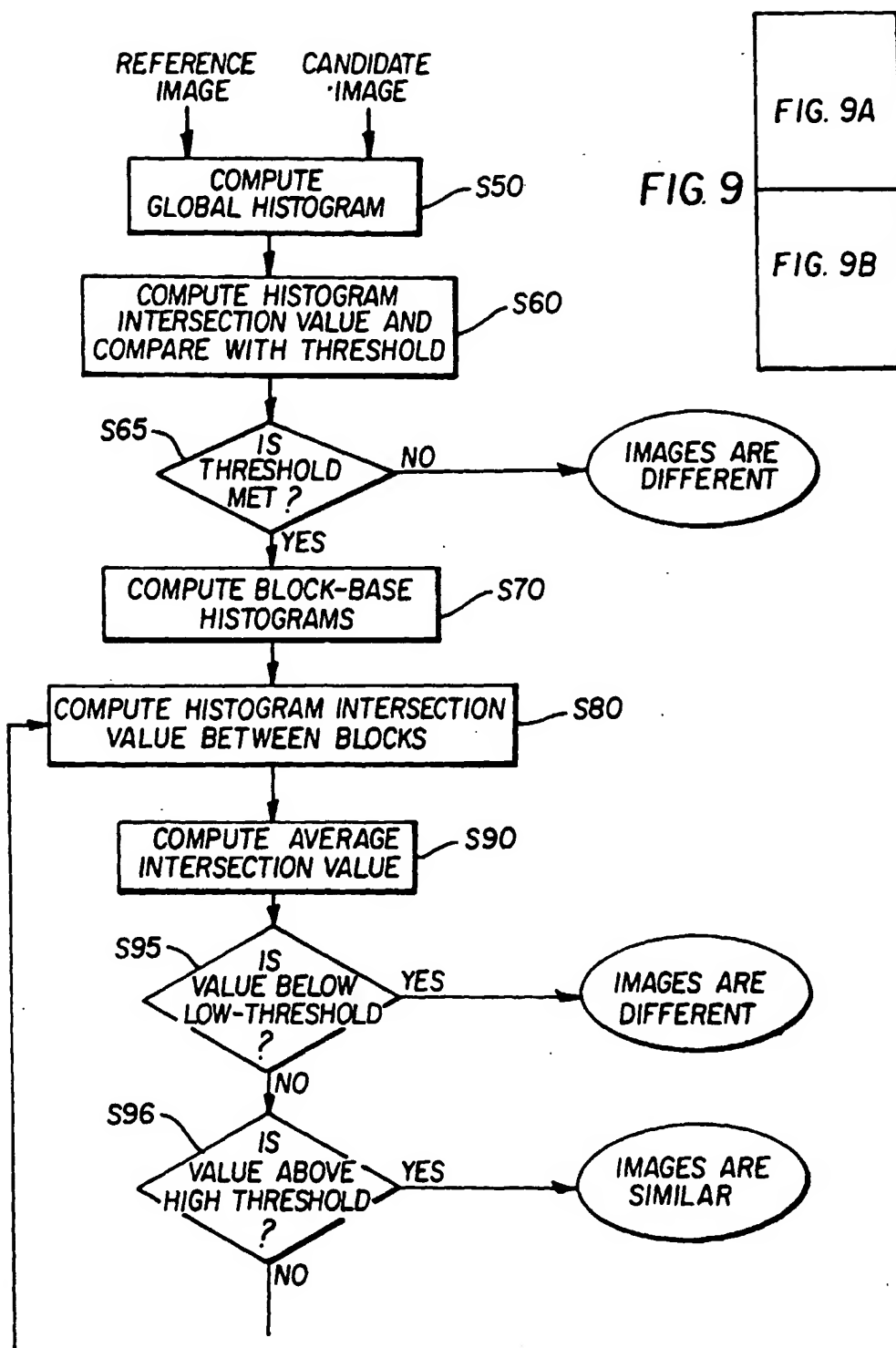


FIG. 9A

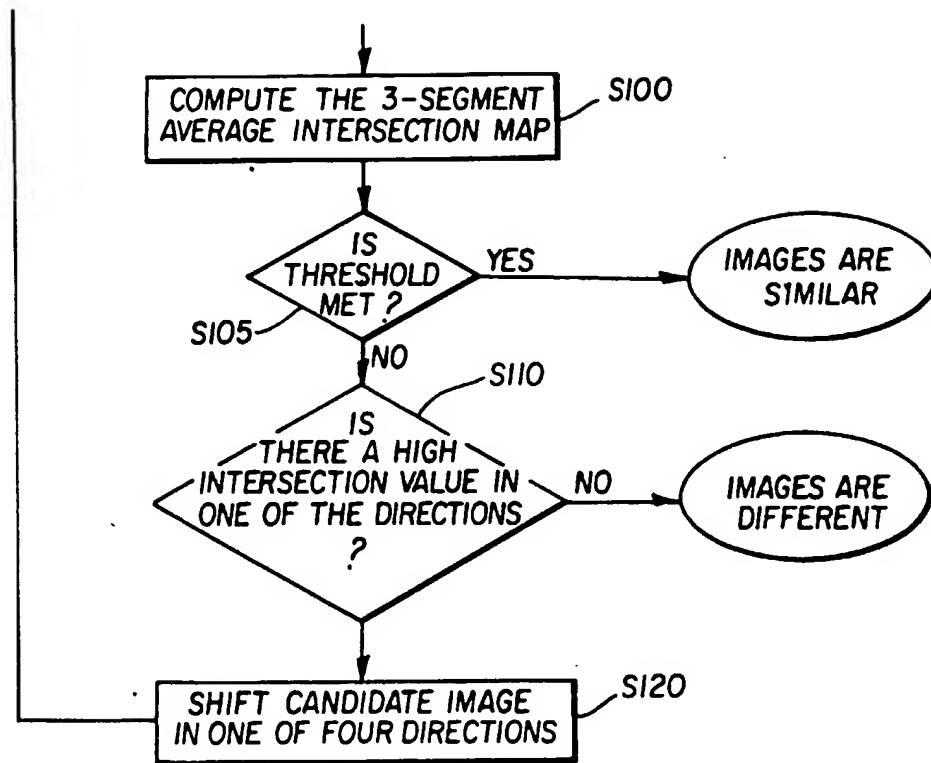


FIG. 9B

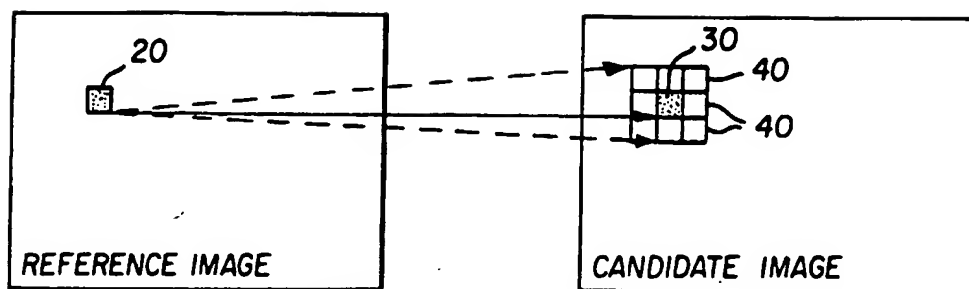


FIG. 10

70a			70b			70c		
0.32	0.45	0.11	0.34	0.56	0.76	0.87	0.24	0.09
0.22	0.10	0.07	0.41	0.48	0.80	0.77	0.44	0.26
0.03	0.16	0.37	0.46	0.68	0.69	0.75	0.21	0.02
0.21	0.34	0.24	0.56	0.87	0.51	0.48	0.11	0.14
0.26	0.22	0.19	0.27	0.38	0.23	0.31	0.12	0.25
0.01	0.08	0.18	0.15	0.18	0.21	0.12	0.20	0.21

FIG. 11

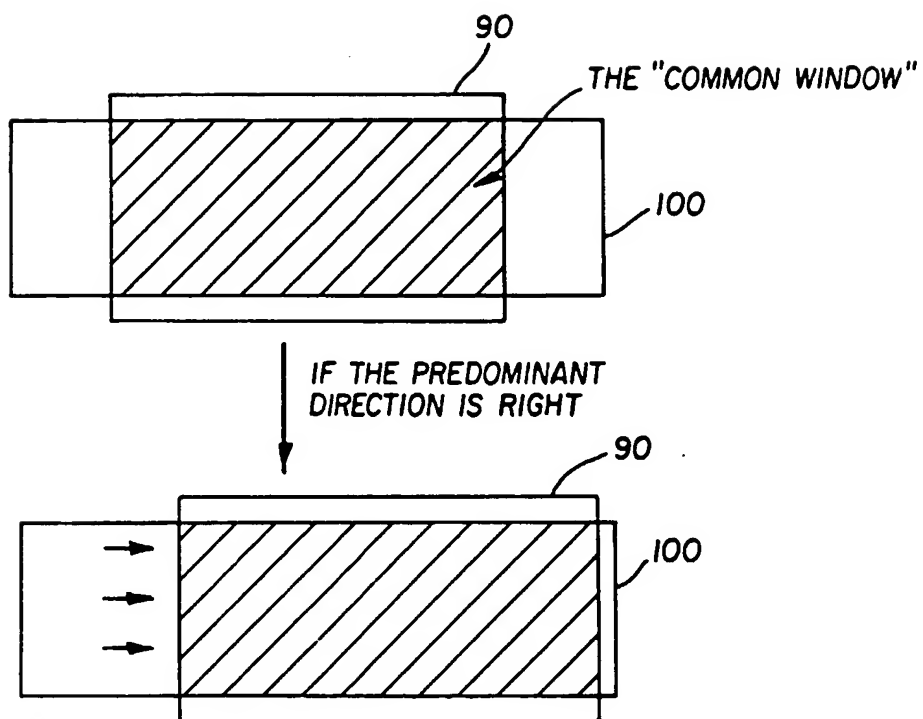


FIG. 12